

## **PROJECT 3: EVALUATION OF IR MODELS**

Mahek Sangwan  
UB Person Number: 50376629  
[maheksan@buffalo.edu](mailto:maheksan@buffalo.edu)

### **1 Overview**

The goal of this project is to implement various IR models, evaluate the IR system and improve the search result based on the understanding of models, the implementation and the evaluation. We are given twitter data in three languages - English, German and Russian, 15 sample queries and the corresponding relevance judgement. We have to index the given twitter data by Solr, implement Vector Space Model and BM25 based on Solr, and evaluate the two sets of results using Trec\_Eval program. Based on the evaluation result, we are asked to improve the performance in terms of the measure Mean Average Precision (MAP).

### **2 Implementation in Solr**

- Two separate cores were created on Solr, for each IR Model. These were named as "IRF20P3\_BM25" and "IRF20P3\_VSM".
- The "managed-schema" files in the conf folder of each core were edited as follows and renamed to "schema.xml"
  - In BM25 the similarity class added was: solr.BM25SimilarityFactory.
  - In VSM the similarity class added was: solr.ClassicSimilarityFactory.

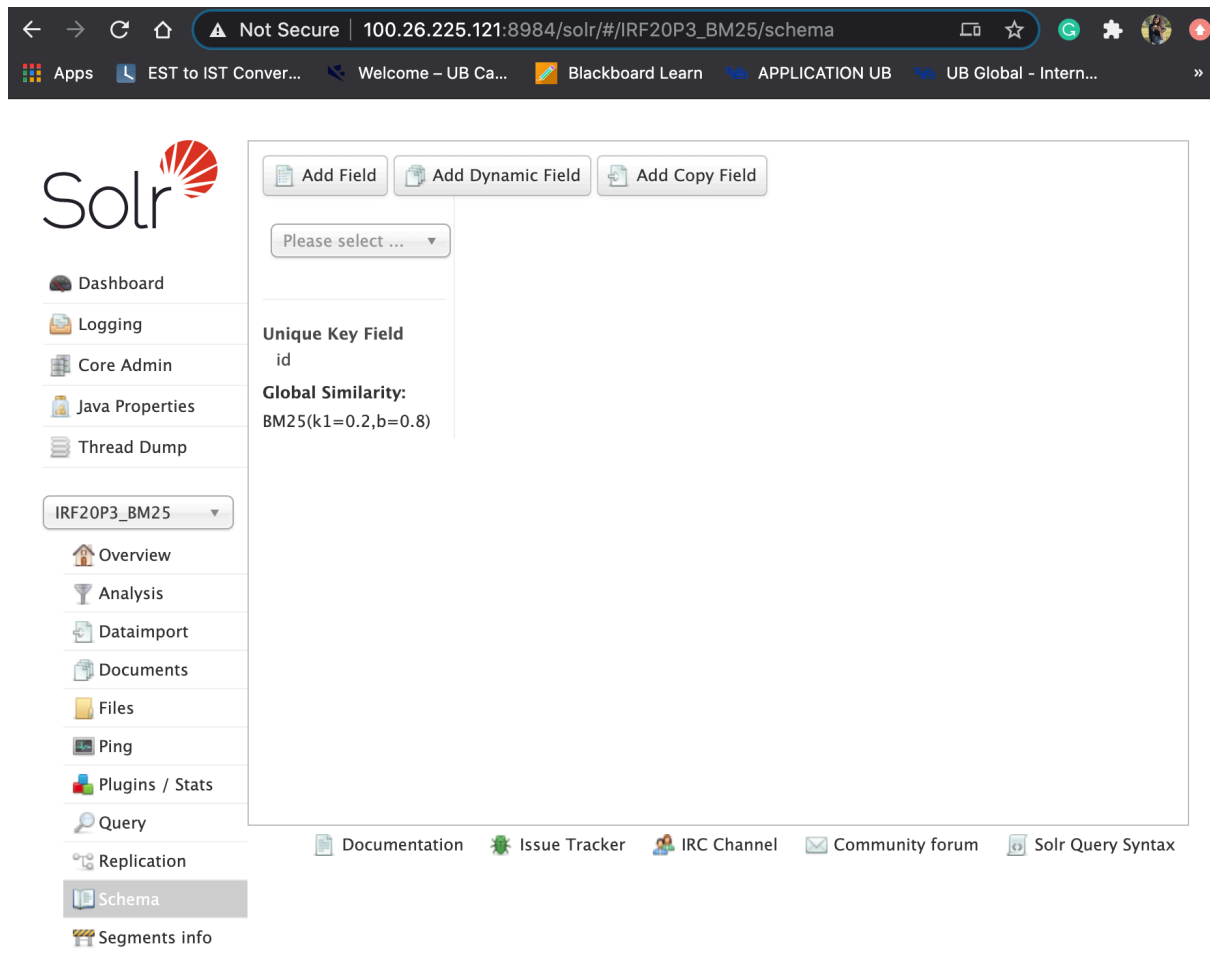
### **3 Obtaining MAP values**

To convert the query results of Solr in a format acceptable for the trec\_eval function, a python script json\_to\_trec.py was used.  
The default values

### **4 Parameter Tuning for BM25**

The parameters b and k1 were tuned to improve the MAP values for the BM25 model. The default value of b= 0.75 and k1= 1.2. Using trial and error these were altered to get a higher accuracy.

b	k1	MAP values
0.75	2	0.6669
0.75	1.5	0.6715
0.5	1.2	0.6697
0.9	1.2	0.6696
1	1.2	0.6674



## 5 Query Parser

Instead of the Standard Query Parser, the Dismax query parser was used as it is more error tolerant and is more apposite for user specific queries. The field types of text\_xx files were converted to their respective \_xx field to provide filters like LowerCaseFilterFactory, PorterStemFilterFactory etc.

For the German queries, we used GermanNormalizationFilterFactory and GermanLightStemFilterFactory as well.

For the Russian queries, we used field boosting to improve the relevance of fields. These were:

- tweet\_hashtags: factor = 3
- text\_en: factor = 6
- text\_de: factor = 6
- text\_ru: factor = 6

This improved the MAP values are as follows:

BM25	0.7073
VSM	0.6862

## 6 Synonyms

We added a list of synonyms for words that were assumed to be occurring often in the queries to improve performance and the relevance. The improved MAP values are as follows:

BM25	0.7266
VSM	0.7193

## 7 Conclusion

Thus, after performing the aforementioned steps the final MAP values are:

1. BM25: 0.7266 (for  $b=0.8$  and  $k_1=0.2$ )
2. VSM: 0.7193